

RESEARCH

Open Access

# No-reference image blur assessment using multiscale gradient

Ming-Jun Chen\* and Alan C Bovik

## Abstract

The increasing number of demanding consumer video applications, as exemplified by cell phone and other low-cost digital cameras, has boosted interest in no-reference objective image and video quality assessment (QA) algorithms. In this paper, we focus on no-reference image and video blur assessment. We consider natural scenes statistics models combined with multi-resolution decomposition methods to extract reliable features for QA. The algorithm is composed of three steps. First, a probabilistic support vector machine (SVM) is applied as a rough image quality evaluator. Then the detail image is used to refine the blur measurements. Finally, the blur information is pooled to predict the blur quality of images. The algorithm is tested on the LIVE Image Quality Database and the Real Blur Image Database; the results show that the algorithm has high correlation with human judgments when assessing blur distortion of images.

**Keywords:** No-reference blur metric, Gradient histogram, Multi-resolution analysis, Information pooling

## 1. Introduction

With the rapid and massive dissemination of digital images and videos, people live in an era replete with digitized visual information. Since many of these images are of low quality, effective systems for automatic image quality differentiation are needed. Although there are a variety of effective full-reference (FR) quality assessment (QA) models, such as the PSNR, the structural similarity (SSIM) index [1,2], the visual information fidelity index [3], and the visual signal-to-noise ratio (VSNR) [4], models for no-reference (NR) QA have not yet achieved performance that is competitive with top performing FR QA models. As such, research in the area of blind or NR QA remains quite vital.

There are many artifacts that may occur in a distorted image, such as blocking, ringing, noise, and blur. Unlike FR QA, where a reference is available to test against any distortion, NR QA approaches generally seek to capture one or a few distortions. Here we are mainly concerned with NR blur assessment, which remains an important problem in many applications. Generally, humans tend to conclude that images with more detail are of higher quality. Of course, the question is not so simple, since

blur can be space-variant, may depend on depth-of-field (hence effect foreground and background objects differently), and may depend on what is being blurred in the image.

A number of NR blur indices have been developed, the majority of which are based on the analyzing luminance edges. For example, the sharpness measurement index proposed by Caviedes and Gurbuz [5] is based on local edge kurtosis. The blur measurement metric proposed by Marziliano et al. [6] is based on analyzing of the width or spread of edges in an image, while their other work is based on an analysis of edges and adjacent regions in an image [7]. Chuang et al. [8] evaluate blur by fitting the image gradient magnitude to a normal distribution, while Karam et al. develop a series of blur metrics based on the different types of analysis applied to edges [9-13].

Other researchers have studied blur assessment by frequency domain analysis of local DCT coefficients [14], and of image wavelet coefficients [15-17]. These methods generally rely on a single feature to accomplish blur assessment. While some of these algorithms deploy simple perceptual models in their design [7,9,11,12,17], a theme that we extend in our approach. Specifically, we use a model of neural pooling of the responses of correlated neuronal populations in the primary visual cortex

\* Correspondence: mjchen@mail.utexas.edu

Department of Electrical & Computer Engineering, Laboratory for Image and Video Engineering, The University of Texas at Austin, Austin, TX, USA

[18]. The use of multiple features combined using machine learning methods has also been studied [19,20].

We are also inspired by recent progress on utilizing natural scene statistics (NSS) to improve image processing algorithms. Natural images obey specific statistical laws that, in principle, might be used to distinguish natural images from artificially distorted images [21]. In this regard, images that are blurred beyond a norm (e.g., more than the band limit provided by the normal human lens at optical center) may measurably depart from statistical “naturalness.” By this philosophy, we may anticipate that NR blur indices can be designed that analysis image statistics. Indeed, Sheikh et al. successfully used NSS for NR QA of JPEG-2000 distorted images [22]. In their work, specific NSS features drawn from the gradient histogram were used.

Here we develop a new blur assessment index that operates in a coarse-to-fine manner. First, a coarse blur measurement using gradient histogram features is deployed that relies on training a probabilistic support vector machine (SVM). A multi-resolution analysis is then used to improve the blur assessment, deploying a model of neural pooling in cortical area V1 [18]. The overall algorithm is shown to agree well with human subjectivity.

The rest of the paper is organized as follows: Section 2 describes the way in which NSS are used. Section 3 describes the coarse-scale NR blur index. Section 4 extends the metric using multi-resolution analysis. Section 5 explains the use of the neural pooling model. The overall NR blur index is evaluated in Section 6, and concluding remarks are given in Section 7.

## 2. Natural image statistics

Recent research on natural image statistics have shown that natural scenes belong to a small set in the space of all possible image signals [23]. One example of a natural scene property is the greater prevalence of strong image gradients along the cardinal (horizontal and vertical) orientations, in images projected from both indoor and outdoor scenes. A number of researchers have developed statistical models that describe generic natural images [19] (including images of man-made scenes).

Although images of real-world scenes vary greatly in their absolute color distributions, image gradients generally have heavy tailed distributions [24]. Natural image gradient magnitudes are mostly small, yet take large values significantly more often than a Gaussian distribution. This corresponds to the intuition that images often contain large sections of smoothly varying intensities, interrupted by occasional abrupt changes at edges or occlusive boundaries. Blurred images do not have sharp edges, so the gradient magnitude distribution should have greater relative mass at small or zero values. By

example, Figure 1 shows a sharp and a blurred image. Figure 2 shows the distribution of their respective gradients.

Liu et al. [25] and Levin [26] have demonstrated that measurements on the heavy tailed distributions of gradients can be used for blur detection. Liu et al. used the gradient histogram span as a feature in their classification model. Levin fits the observed gradient histogram using a mixture model.

## 3. Probabilistic SVM For blur assessment

Based on our discussion of NSS, we seek to evaluate the distance between the gradient statistics of an (ostensibly distorted) image and a statistical model of natural scenes. This distance can then be used for image QA.

A classification method is used to measure the distance. We classify the images into two groups. One is tagged as “sharp” and the other as “blurred.” Using the probabilistic SVM classification model, confidence values are computed that represent the distance between the test image and the training set. A higher confidence value implies a higher certainty of the classification result. In this case, this means that the test sample is closer to the assigned class center, i.e., the statistic of the test image is closer to that of “sharp” or “blurred” images.

We chose to use a SVM [27] as our classification model. The main reason for using SVM is that it works well for classifying a few classes with few training samples. This is highly suitable for our application having only two classes. Moreover, SVM allows substitution of kernels to achieve better classification results. Although here we only use the default kernel, the possibility of modifying the kernel leaves room for performance improvement.

Due to the limited scope of the coarse evaluation of the image, we use the entire gradient histogram as a feature, rather than simple measured parameter such as the mean or the slope of the histogram [25,26]. While this implies a fairly large number of features, it is not very large, and the small number of classes ensures reasonable computation. We describe the training procedure and the dataset used in Section 6.

After applying probabilistic SVM classification on an image, a label that indicates its class and a confidence score that indicates the degree of confidence in the decision are obtained. Then the coarse quality score of the image is defined simply as:

$$QS - SVM(x) = \begin{cases} 50 + 50 \cdot \text{confidence}, & \text{if } x \text{ is classified as sharp} \\ 50 \cdot (1 - \text{confidence}), & \text{if } x \text{ is classified as blurred} \end{cases} \quad (1)$$

## 4. Multi-resolution NR QA of blur

As in most other areas of image processing and analysis, multi-resolution methods have afforded improved



**Figure 1 Exemplar images.** Left: blurred image A. Right: Bottom: Sharp image B.

performance relative to single-resolution methods for image QA [2,4]. In the following, we modify QS-SVM using information derived from a multi-resolution decomposition.

Applying a wavelet decomposition on an image is a natural way to examine local spatio-spectral properties that may reveal whether the image has been modified. For example, Figure 3 shows a sharp natural image decomposed by a two-level wavelet, while Figure 4 shows the decomposed blurred image. The sharp image is a high-quality image from the LIVE database. The blurred image was modified by a Gaussian low-pass filter. We used the 2D analysis filter bank developed in [28] to analyze the image. From Figures 3 and 4, it is apparent that the sharp image contains significant horizontal and vertical energy in the high bands, while the blurred image does not. As a simple measure of sharpness, we sum the horizontal and vertical responses in the high band to produce a *detail map*. Figure 5 shows the detail map derived from the sharp image in Figure 3.

A quality (or sharpness) score that combines QS-SVM with multi-resolution analysis follows:

$$\text{Blur quality score} = (\text{QS-SVM})^{r_0} \prod_{i=1}^N (\text{DS}_i)^{r_i} \quad (2)$$

where  $N$  is the number of layers in the wavelet decomposition, and QS-SVM is the score obtained by analyzing the original image using the probabilistic SVM model described in the preceding. Further,  $\text{DS}_i$  is the *detail score* obtained from the detail map of layer  $i$ . The detail score at wavelet level  $i$  is defined:

$$\text{DS}_i = \frac{\sum_{m=1}^{W_i} \sum_{n=1}^{H_i} \nabla_i(m, n)}{W_i \cdot H_i} \quad (3)$$

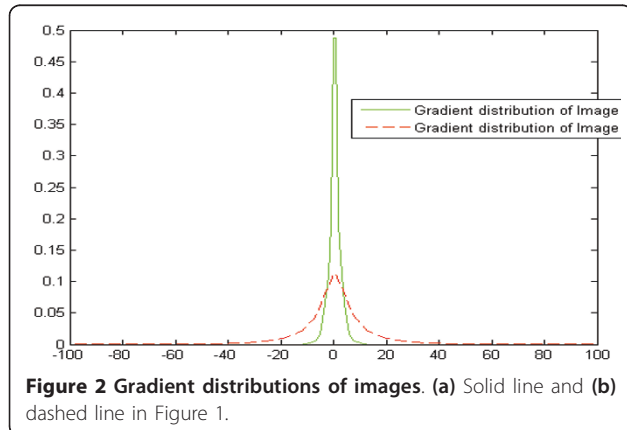
where  $W_i$  and  $H_i$  are the pixel dimensions of the sub-band image that  $\text{DS}_i$  is defined on, and  $\nabla_i(m, n)$  is the gradient magnitude value of the subband image at coordinate  $(m, n)$ .

Blur quality score is the final blur evaluation result, which is the weighted (by exponents) product of the full-resolution score QS-SVM and the values of DS from each layer. The parameters  $r_i$  are normalized exponents:  $\sum_{i=0}^N r_i = 1$ .

## 5. Decoding the neural responses

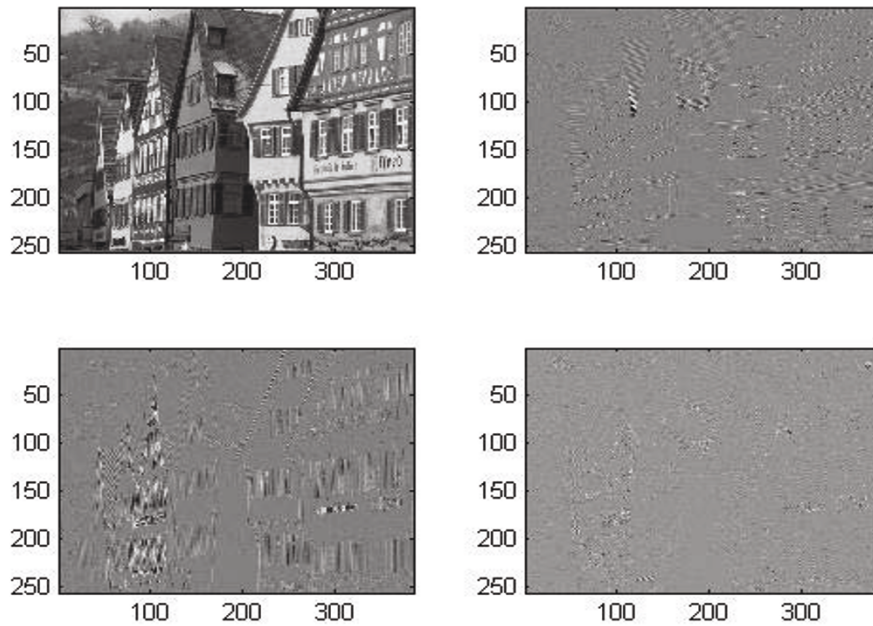
Perceptual models have played an important role in the development of image QA algorithms. These have included image masking models [29], cortical decompositions [30], extra-cortical motion processing [31], and foveation [29,32-34], among others.

The visual model we will use is based on foveal (non-peripheral) processing of visual information. The central two degrees of high-resolution imagery that is projected onto the human fovea subtends roughly equivalent twice the width of a thumbnail at arm's length [35]. In [9], a viewing model is used to derive the use of  $64 \times 64$  blocks to approximate the size of image patches projected from display onto the fovea (see Figure 6). In this



**Figure 2 Gradient distributions of images.** (a) Solid line and (b) dashed line in Figure 1.

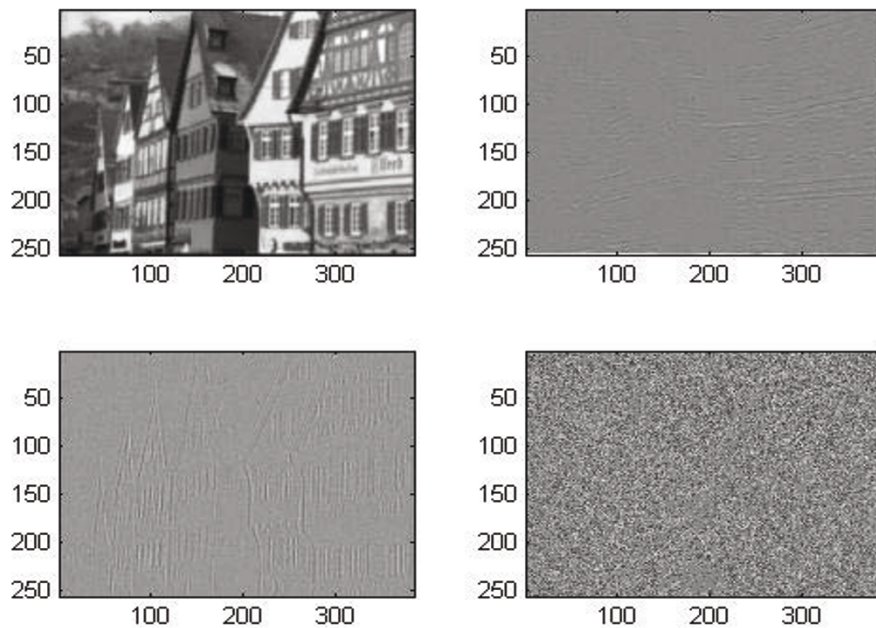




**Figure 3 Wavelet decomposition of natural image.** Top left: low band response. Top right: horizontal high band response. Bottom left: vertical high band response. Bottom right: high band response.

viewing model, a subject is assumed to be sitting in front of a 24"  $\times$  18" LCD screen with a resolution of 1680  $\times$  1050 pixels. The width of foveal vision at arm's length is assumed to be about 1.2", while the viewing distance is assumed to fall in the range 36-54"

(approximately 2-3 times the screen height). The arm length of the viewer is assumed to be 33 in.. Then, the width of span of foveal vision on the screen falls between 76 ( $1050/18 \times 1.31$ ) and 116 ( $1050/18 \times 2$ ) pixels.



**Figure 4 Wavelet decomposition of blurred image.** Top left: low band response. Top right: horizontal high band response. Bottom left: vertical high band response. Bottom right: high band response.



**Figure 5** Detail map computed from image in Figure 3.

Since a block size of  $2''$  along each dimension facilitates optimization and allows better memory management (aligned memory allocation), the choice of a  $64 \times 64$  block size is a decent approximation. We then apply the blur QA method described in Section 4 on each of these blocks.

When a human observer studies an image, thus arriving at a sense of its quality, she/he engages in a process of eye movement, where visual saccades place fixations at discrete points on the image. Image quality is largely

decided by information that is collected from these foveal regions, with perhaps, additional information drawn from extra-foveal information. The overall perception of quality drawn from these fixated regions might be described as “attentional pooling,” by analogy with the aggregation of information from spatially distributed neurons. We utilize the results of a study conducted by Chen et al. [18] to formulate such an attentional pooling strategy. In this study, the authors examined the efficacy of different patch pooling strategies in a primate undergoing a visual (Gabor) target detection task.

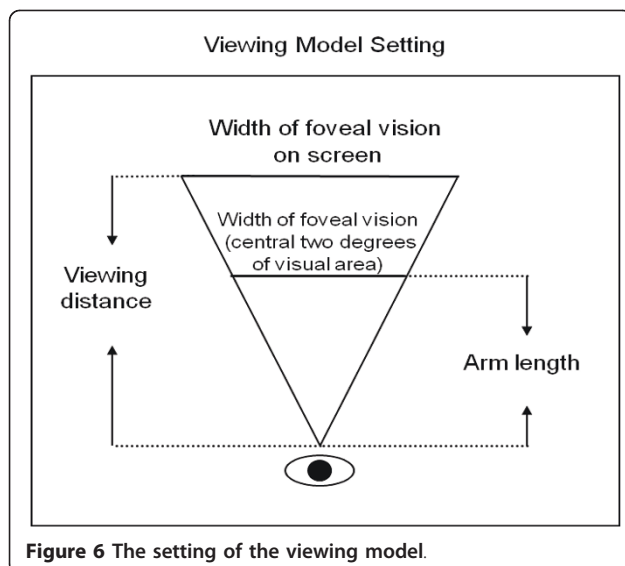
The authors of [18] used voltage sensitive dyed images to measure the population responses in primary visual cortex of monkeys performing a demanding visual target detection task. Then, they evaluated the effects of different decoding strategies in predicting the target pattern from measured neural responses in primary visual cortex. The pooling process they considered used a linear summation model:

$$X_{\text{pooled}} = \sum_{i=1}^n w_i x_i \quad (4)$$

where  $w_i$  is the weight applied to the neuronal amplitude response  $x_i$ .

The pooling rules they studied are as follows:

1. Maximum average amplitude:  $w_i \neq 0$  only for the patch having maximum average neuronal response amplitude.



**Figure 6** The setting of the viewing model.

2. Maximum  $d'$ :  $w_i \neq 0$  only for the patch having maximum  $d'$
3. Maximum amplitude:  $w_i \neq 0$  only for the site with maximum amplitude in a given trial
4. Mean amplitude:  $w_i = 1/n$
5. Weighted average amplitude:  $w_i$  is proportional to the average amplitude response of  $x_i$
6. Weighted  $d'$ :  $w_i$  is proportional to  $d'$
7. Optimal

where  $d'$  is the SNR of the neuronal responses across trials:

$$d' = |E_S - E_N| / \sqrt{\frac{\sigma_S^2 + \sigma_N^2}{2}} \quad (5)$$

where  $E_S$  is the mean response amplitude in target present trials (signal trials),  $E_N$  is the mean amplitude of the response in target-absent trials (noise trials) and  $\sigma_S$  and  $\sigma_N$  are the corresponding standard deviations. The “optimal” pooling 7 is obtained under the assumption that the neuronal response at each site is Gaussian distributed and independent across trials (although not across space and time within a trial). The optimal set of weights is defined as the product of the inverse of the response covariance matrix the vector of mean differences in response between the signal and noise trials. Their experimental result is shown in Figure 7.

From Figure 7, we can see that the maximal average pooling rules (Rules 1 and 2) perform better than the trial maximum (Rule 3), average pooling rules (Rule 4) and weighted pooling rules (Rules 5 and 6). When applying analogous pooling rules to the image blur assessment problem, we observe that since distinct signal and noise trials do not exist in our case (and in any

case the Gaussian assumption is questionable), so we cannot apply the optimal pooling rule (Rule 7). Further, the SNR  $d'$  is not available as required by Rules 2 and 6. Hence, we choose the maximum average amplitude as our pooling rule. The slight difference here is that with a single (algorithmic) “trial” an average amplitude value is not available, while the maximum amplitude (Rule 3) is unreliable. Instead, we use the average of the maximum  $p\%$  of responses as a pooling strategy. The pooling strategy was applied only on activated neurons; hence we applied the pooling only on activated blocks, where a block was taken to be activated if the mean of the luminance values in the block is higher than 20. Therefore, the final blur quality score is calculated as

$$\text{Blur quality score} = (\text{QS} - \text{SVM})^{r0} * \prod_{i=1}^N (\text{Pool}(\text{DS}_i))^{r1} \quad (6)$$

where

$$\text{Pool}(\text{DS}_i) = \sum_{k=1}^n w_{ki} \text{DS}_{ki} \quad (7)$$

where  $\text{DS}_{ki}$  is the detail response of block  $k$  from layer  $i$ , and  $w_{ki} = 1/p$  if the detail responses of block  $k$  in layer  $i$  belong to the largest 10% of detail responses of all activated blocks in the layer; otherwise  $w_{ki} = 0$ . Here,  $p$  is nominally set to 10. The blocking analysis and pooling are only applied on the multi-resolution part, since the NSS mentioned in Section 2 are based on the statistics of whole images.

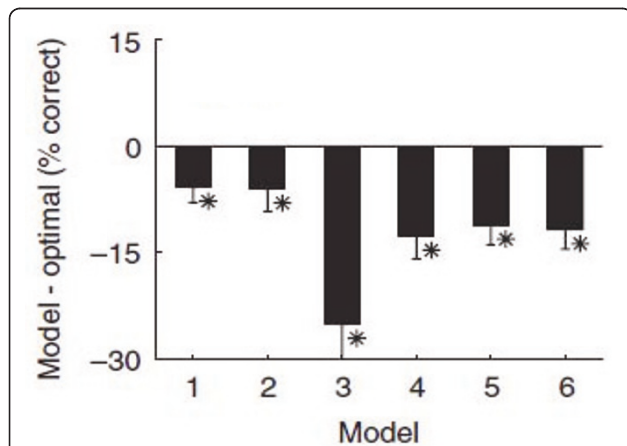
## 6. Experiments and results

The LIVE image quality database [36] and the real blur image Database [37] were used to evaluate the performance of our algorithm. The experiments in Sections 6.1-6.3 were conducted on the LIVE database to gain insights into the performance of algorithms that combine different blur assessment factors. The performances are also compared to the performance of multi-scale SSIM (or MS-SSIM, a popular and effect FR QA method).

Then in Section 6.4, the Real blur database (586 Images) is used as a challenging test by which we compare our results with other NR QA blur metrics. The LIVE image database includes DMOS subjective scores for each image and several types of distortions. The experiment was performed only on the blur images (174 images). All of the images in the LIVE database are blurred globally. Samples of these images are shown in Figure 8. A total of 760 images were used for testing.

### 6.1. Performance of SVM Classification

To train the coarse SVM classifier, we used 240 training samples which were marked as “sharp” or “blurred.” The



**Figure 7** Comparison of detection sensitivity of candidate pooling rules. Asterisks indicate rules with performance significantly different from the optimal (bootstrap test,  $p < 0.05$ ).





**Figure 8** Sample images from the LIVE image quality database. From top-left to bottom-right, increasing Gaussian blur is applied.

training samples were randomly chosen and some of them are out-of-focus images. Due to the unbalanced quality of the natural training samples (there were more sharp images than naturally blurred images), we applied a gaussian blur to some of the sharp samples to generate additional blurred samples. The final training set included 125 sharp samples and 115 blurred samples. The training and test sets do not share content.

When tagging samples, if an original image's quality was mediocre, the image was duplicated; one copy marked as "blurred" and the other marked as "sharp," with both images used for training. This procedure prevents misclassifications arising from marking mediocre image as "sharp" or "blurred." This duplication was applied to lower the confidence when classifying mediocre samples.

Note that DMOS scores of these images we are not required to train the SVM. Images were simply tagged as "blurred" or "sharp" to train the SVM. Likewise, the output of the probabilistic SVM model is a class type ("blurred" or "sharp") and a confidence level. The class type and confidence level are used to predict the image quality score.

The algorithm was evaluated against the LIVE DMOS scores using the Spearman rank order correlation coefficient (SROCC). The results are shown in Table 1.

In Table 1, QS-SVM means blind blur QA using probabilistic SVM, PSNR means peak signal to noise ratio, and MS-SSIM means multi-scale structure similarity index. To obtain an objective evaluation result, we compared our method to FR methods tested on the same database as in [4,38].

As can be seen, the coarse algorithm QS-SVM delivered lower SROCC scores than the FR indices, although the results are promising. Of course, QS-SVM is not trained on DMOS scores, hence does not fully capture the perceptual elements of blur assessment.

## 6.2. Performance with multi-resolution decomposition

We began by estimating which layers of the wavelet decomposition achieve the best QA result on the LIVE database. We found the correlations between the DS

**Table 1** Comparison of the performance of VQA algorithms

Prediction model	SROCC
QS-SVM	0.6136
PSNR (FR)	0.7729
VSNR (FR)	0.932
MS-SSIM (FR)	0.9425

**Table 2 QA performance using different layers**

Prediction model	SROCC
QS-SVM	0.6136
DS <sub>0</sub>	0.6583
DS <sub>1</sub>	0.8884
DS <sub>2</sub>	0.7733
DS <sub>3</sub>	0.5587

scores and human subjectivity for each layer. The performance numbers are shown in Table 2.

In Table 2, DS<sub>0</sub> is the detail score computed from the original image. The experiment shows the SROCC score of DS<sub>1</sub> to be significantly higher than for the other layers. The detail map at this middle scale appears to deliver a high correlation with human impression of image quality.

Next we combined the QA measurement in different layers, omitting level 3 because of its poor performance. Table 3 shows the results of several combinations of algorithms. The parameters  $r_i$  of each combination were determined by regression on the training samples.

Table 3 shows that, except for combination with QS-SVM, all other combinations with DS<sub>1</sub> did not achieve higher performance than using only DS<sub>1</sub>. This result is consistent with our other work in FR QA, where we have found that mid-band QA scores tend to score higher than low-band or high-band scores. Adding more layers did not improve performance here. The highest performance occurs by combining DS<sub>1</sub> with QS-SVM ( $r_0 = 0.610$ ,  $r_1 = 0.390$ ), yielding an impressive SROCC score of 0.9105. Combination QS-SVM with DS<sub>2</sub> ( $r_0 = 0.683$ ,  $r_2 = 0.317$ ) also improved performance relative to DS<sub>2</sub>, suggesting that QS-SVM and the DS scores offer complementary measurements.

### 6.3. Performance with pooling strategy

We studied the performance of different pooling rules in our system. The system is described by (6), using maximum  $p\%$  pooling, average pooling (Rule 4 in Section 5), and weighted pooling (Rule 5 in Section 5), applied to QS-SVM-DS<sub>1</sub>. Using tenfold cross-validation with fixed parameters  $r_0 = 0.610$  and  $r_1 = 0.390$ , the performance attained is given in Table 4. Table 4 shows that the

**Table 3 QA performance using different combinations of layers**

Prediction model	SROCC
DS <sub>0</sub> -DS <sub>1</sub>	0.8884
DS <sub>1</sub> -DS <sub>2</sub>	0.8884
QS-SVM-DS <sub>1</sub>	0.9105
QS-SVM-DS <sub>1</sub> -DS <sub>2</sub>	0.9105
QS-SVM-DS <sub>2</sub>	0.8428

**Table 4 QA performance numbers by tenfold cross-validation**

Pooling rule	SROCC
Maximum $p\%$	0.9248
Average	0.9004
Weighted	0.9080

Different pooling rules were applied on the blurred image portion of the LIVE Image Quality Database

performance of using different pooling rules in our system is consistent with the results found in [18]. The maximum  $p\%$  pooling method improves the performance (the SROCC score is increased from 0.9004 to 0.9248).

All parameters in our system were kept fixed ( $p = 10$ ,  $r_0 = 0.610$  and  $r_1 = 0.390$ ) to enable fair comparisons with other algorithms. The number  $p$  came from cross-validation across two databases. Table 5 illustrates the final performance of our algorithm as compared to other NR and FR blur metrics. The performance of our algorithm is better than PSNR and very close to CPBD [10] and to FR QA models when conducted on the blurred image portion of the LIVE Image Quality Database. The plot of predicted objective quality (following logistic regression) against DMOS scores from the LIVE Image Quality Database is shown in Figure 9.

### 6.4. Challenging blur database

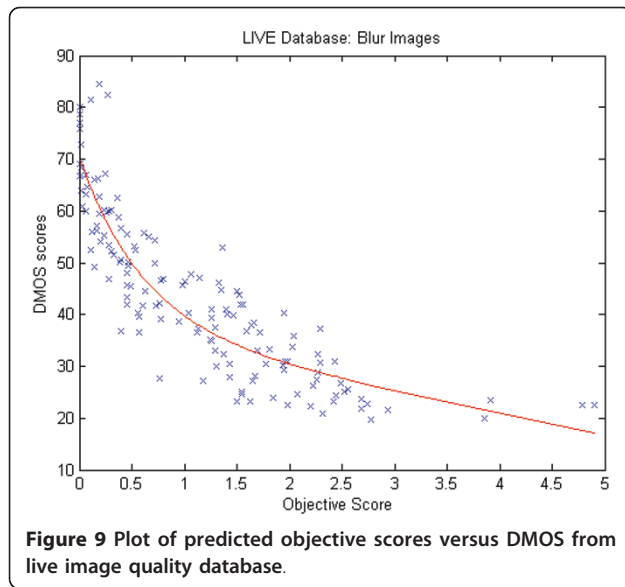
Our foregoing experiments on the LIVE database were by way of algorithm design and tuning, and not performance verification. To verify the performance of our algorithm, we conducted an experiment on a real blurred image database. The database contains 585 images with resolutions ranging from  $1280 \times 960$  to  $2272 \times 1704$  pixels.

The images in this database were taken by consumer cameras and are classified into five classes as “Unblurred” (204 images), “Out-of-focus” (142 images), “Simple Motion” (57 images), “Complex Motion” (63 images) and “Other” (119 images). The images in the

**Table 5 Summary of QA performance of different algorithms on the blurred image portion of the LIVE Image Quality Database**

Prediction model	SROCC
QS-SVM	0.6136
PSNR (FR)	0.7729
QS-SVM-DS <sub>1</sub>	0.9105
QS-SVM-Pool(DS <sub>1</sub> )	0.9352
VSNR (FR)	0.932
MS-SSIM (FR)	0.9425
CPBD	0.9430



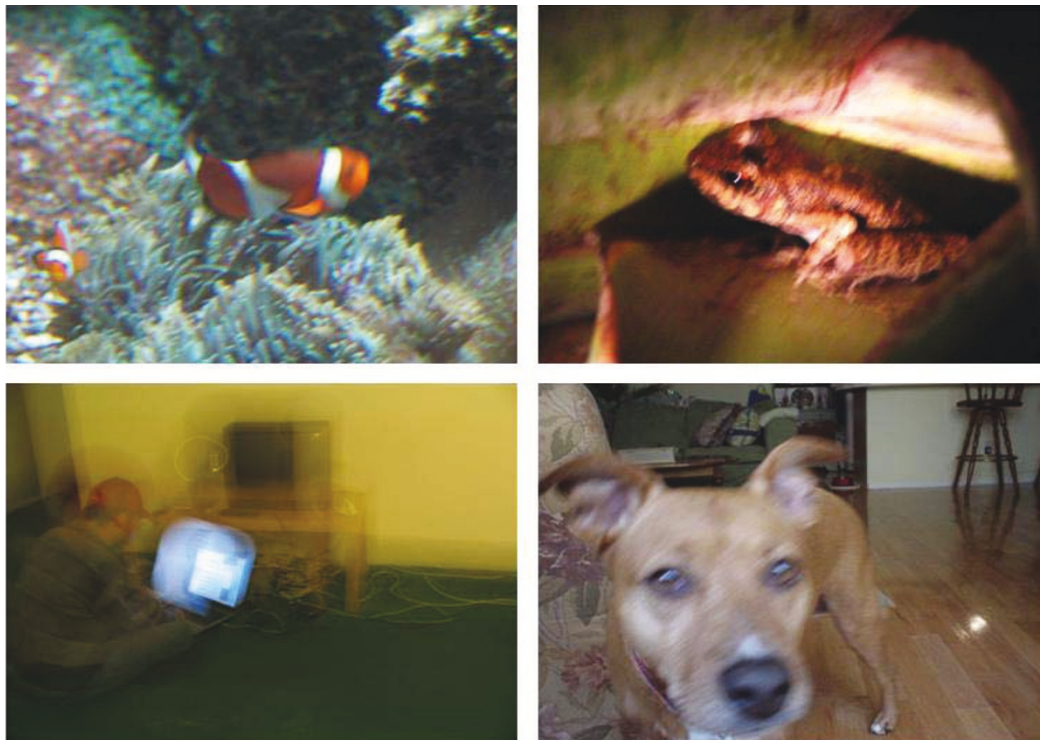


“Out-of-focus” (142 images), “Simple Motion” (57 images), “Complex Motion” (63 images) and “Other” (119 images). The images in the “Out-of-focus” class are global out-of-focus images. The “Simple Motion” class has images that are blurred because of close-to-linear camera movements and the “Complex Motion” class has

images which are blurred because of the more complex motion paths. Finally, the “Other” class includes any other types of degradation. It may include any combination of the main classes. For instance, the image with localized out-of-focus blur (mixed “Unblurred” and “Out-of-focus”) was classified into the “Other” class. Sample images are shown in Figure 10. The raw MOS scores of the database are provided. We eliminated 20% (maximum 10% and minimum 10%) of the grades on each image as outliers, so that the average (trimmed mean) of the 80% grades was used as the MOS score of each image.

We used tenfold cross-validation and report the SROCC numbers from applying several different pooling rules. As shown in Table 6, the maximum  $p\%$  pooling method yields the best performance (0.5858). Although the improvement is not significantly large, this method showed the best performance on both databases.

By examining the experimental results from the LIVE Image Quality Database and the Real Blur Image Database, we found that there is a significant performance difference of the models on these two databases. The LIVE database includes synthetically and globally blurred sample images. The task of performing QA on a globally blurred image is less complex and harder to relate to perceptual models. On LIVE, our proposed



**Figure 10** Sample images from the real blur database. Top left: Out-of-focus image. Top right: Simple motion blur. Bottom left: Complex motion blur. Bottom right: Others (partial blur case).

**Table 6 Blur QA performance of applying different pooling rules on real blur database**

Pooling rule	SROCC
Maximum $p\%$	0.5858
Average	0.5604
Weighted	0.5542

method of pooling showed significant improvement (from 0.9 to 0.925). However, on the Real Blur Database, where the blurs are more complex, possibly nonlinear, and spatially variant, blur perception is more complex and probably more correlated with content (e.g., what is blurred in the image?). By example, in the partially blurred image shown in Figure 10 (bottom right), the rating is likely highly affected by image content, object positioning, probable viewer fixation, and so on.

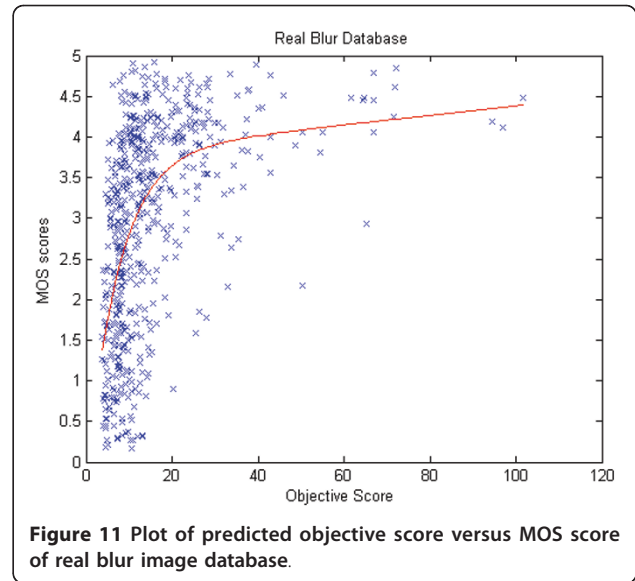
When comparing the performance of our proposed algorithm with other blur assessment algorithms, we refer to the work conducted by Ciancio et al. [20]. In this work, they provided performance levels several algorithms, including a frequency-domain blur index [14], a wavelet-based blur index [15], a perceptually motivated blur index [7], a blur index using a human visual system (HVS) model [11], a local phase coherence blur metric [16], and their own Multi-Features Neural Network Classifier (MFNNC) blur metric [20]. The performance of CPBD [10] is also included. The performance results are shown in Table 7.

Table 7 shows that our proposed blur QA model delivers the best performance amongst the algorithms compared. Although the improvement does not achieve statistical significance as compared with other top-performing models, it consistently shows better performance across a large number of images and across databases. A scatter plot of the scores delivered by our model (following logistic regression) against the MOS scores from the Real Blur Database is shown in Figure 11 showing very good general agreement. Many of the images, such as Figure 12, contain difficult high level

**Table 7 Blur QA performance of different algorithms on real blur database**

Algorithm	SROCC
Frequency domain metric*	0.494
Wavelet-based metric*	0.524
Perceptual metric*	0.336
HVS based metric*	0.474
Local phase coherence metric*	0.523
MFNNC metric*	0.564
Proposed algorithm	0.586
CPBD	0.501

Algorithms marked by asterisk indicates their performance was reported in [20]



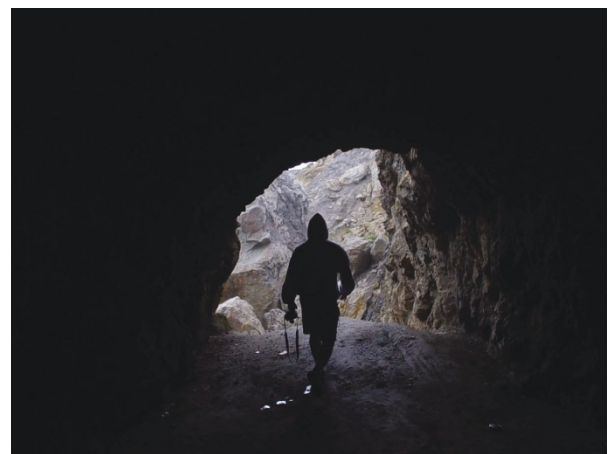
**Figure 11 Plot of predicted objective score versus MOS score of real blur image database.**

content whose interpretation may depend on the observers' preferences regarding composition (and that of the photographer).

## 7. Conclusion

The main contributions of this work are as follows. First, we found that the statistics of the image gradient histogram and a detail map from the image wavelet decomposition can be combined to yield good NR blur QA performance. Second, our results discuss that a perceptually motivated pooling strategy can be used to improve the NR blur index on assessing the blur images.

Performance was demonstrated on the LIVE Image Quality Database and the Real Blur Image Database. As



**Figure 12 Subjects give higher quality to this image MOS is 3.98 (scale from 0 (worst) to 5 (best)), but our algorithm gives low objective score to this image.**

compared with other NR blur metrics, our method yields competitive performance with reasonable complexity.

#### Abbreviations

QA: quality assessment; SVM: support vector machine; FR: full-reference; SSIM: structural similarity; VSNR: visual signal-to-noise ratio; NSS: natural scene statistics; SROCC: Spearman rank order correlation coefficient; HVS: human visual system; MFNNC: multi-features neural network classifier.

#### Competing interests

The authors declare that they have no competing interests.

Received: 15 September 2010 Accepted: 19 July 2011

Published: 19 July 2011

#### References

- Wang Z, Bovik AC, Sheikh HR, Simoncelli EP: **Image quality assessment: from error visibility to structural similarity.** *IEEE Trans Image Process* 2004, **13**(4):600-612.
- Wang Z, Simoncelli EP, Bovik AC: **Multi-scale structural similarity for image quality assessment.** *IEEE Asilomar Conf. Signals, Systems, and Computers* 2003, 1398-1402.
- Sheikh HR, Bovik AC: **Image information and visual quality.** *IEEE Trans Image Process* 2006, **15**(2):430-444.
- Chandler DM, Hemami SS: **VSNR: a wavelet-based visual signal-to-noise ratio for natural images.** *IEEE Trans Image Process* 2007, **16**(9):2284-2298.
- Caviedes JE, Gurbuz S: **No-reference sharpness metric based on local edge kurtosis.** *IEEE International Conference on Image Processing, Rochester, NY* 2002.
- Marziliano P, Dufaux F, Winkler S, Ebrahimi T: **A no-reference perceptual blur metric.** *International Conference on Image Processing, Rochester, NY* 2002.
- Marziliano P, Dufaux F, Winkler S, Ebrahimi T: **Perceptual blur and ringing metrics: application to JPEG2000.** *Signal Process Image Commun* 2004, **19**(2):163-172.
- Chung Y, Wang J, Bailey R, Chen S, Chang S: **A nonparametric blur measure based on edge analysis for image processing applications.** *IEEE Conf Cybern Intell Syst* 2004, 1:356-360.
- Narvekar ND, Karam LJ: **A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection.** *International Workshop on Quality of Multimedia Experience* 2009.
- Narvekar ND, Karam LJ: **A no-reference image blur metric based on the cumulative probability of blur detection (CPBD).** *IEEE Trans Image Process* 2011.
- Ferzli R, Karam LJ: **A human visual system based no-reference objective image sharpness metric.** *IEEE International Conference on Image Processing, Atlanta, GA* 2006.
- Ferzli R, Karam LJ: **A no-reference objective image sharpness metric based on just-noticeable blur and probability summation.** *IEEE International Conference on Image Processing, San Antonio, TX* 2007.
- Varadarajan S, Karam LJ: **An improved perception-based no-reference objective image sharpness metric using iterative edge refinement.** *IEEE International Conference on Image Processing, Chicago, IL* 1998.
- Marichal X, Ma WY, Zhang H: **Blur determination in the compressed domain using DCT information.** *IEEE International Conference on Image Processing, Kobe, Japan* 1999.
- Tong H, Li M, Zhang H, Zhang C: **Blur detection for digital images using wavelet transform.** *IEEE International Conference on Multimedia and EXPO* 2004, 1:17-20.
- Ciancio A, Targino AN, da Silva EAB, Said A, Obrador P, Samadani R: **Objective no-reference image quality metric based on local phase coherence.** *IET Electron Lett* 2009, **45**(23):1162-1163.
- Wang Z, Simoncelli EP: **Local phase coherence and the perception of blur.** *Advances in Neural Information Processing Systems* MIT Press, Cambridge; 2004, 786-792.
- Chen Y, Geisler WS, Seidemann E: **Optimal decoding of correlated neural population responses in the primate visual cortex.** *Nat. Neurosci* 2006, **9**:1412-1420.
- Narwaria M, Lin W: **Objective image quality assessment based on support vector regression.** *IEEE Trans Neural Netw* 2010, **21**(3):515-519.
- Ciancio A, Targino da Costa ALN, da Silva EAB, Said A, Samadani R, Obrador P: **No-reference blur assessment of digital pictures based on multi-feature classifiers.** *IEEE Trans Image Process* 2011, **20**(1):64-75.
- Simoncelli EP: **Statistical models for images: compression, restoration and synthesis.** *Proceeding of the IEEE Asilomar Conference on Signals, Systems, and Computers* 1997.
- Sheikh HR, Bovik AC, Cormack LK: **No-reference quality assessment using natural scene statistics: JPEG2000.** *IEEE Trans Image Process* 2005, **14**(11):1918-1927.
- Ruderman DL: **The statistics of natural images.** *Netw Comput Neural Syst* 1994, **5**(4):517-548.
- Field D: **What is the goal of sensory coding?** *Neural Comput* 1994, **6**:559-601.
- Liu R, Li Z, Jia J: **Image partial blur detection and classification.** *IEEE International Conference on Computer Vision Pattern Recognition* 2008, 1-8.
- Levin A: **Blind motion deblurring using image statistics.** *Neural Information Processing Systems (NIPS)* 2006, 841-848.
- LIBSVM:[http://www.csie.ntu.edu.tw/~cjlin/libsvm/].
- Abdelnour AF, Selesnick IW: **Nearly symmetric orthogonal wavelet bases.** *IEEE International Conference on Acoustic, Speech, Signal Processing (ICASSP)* 2001.
- Teo PC, Heeger DJ: **Perceptual image distortion.** *IEEE International Conference on Image Processing, Austin, TX* 1994.
- Taylor CC, Pizlo Z, Allebach JP, Bouman CA: **Image quality assessment with a Gabor pyramid model of the human visual system.** *SPIE Conference on Human Vision and Electronic Imaging, San Jose, CA* 1997.
- Seshadrinathan K, Bovik AC: **Motion tuned spatio-temporal quality assessment of natural videos.** *IEEE Trans Image Process* 2010, **19**(2):335-350.
- Lee S, Pattichis MS, Bovik AC: **Foveated video quality assessment.** *IEEE Trans Multimedia* 2002, **4**(1):129-132.
- Wang Z, Bovik AC, Lu L, Kouloheris J: **Foveated wavelet image quality index.** *SPIE's 46th Annual Meeting, Proceedings of SPIE, Application of digital image processing* 2001, XXIV:4472.
- Cormack LK: **Computational models of early human vision.** In *The Handbook of Image and Video Processing*. Edited by: Bovik AC. Academic Press; 2000.
- Mark F: *Color Appearance Models* Addison-Wesley, Boston; 1998, 7.
- Sheikh HR, Wang Z, Cormack LK, Bovik AC: **LIVE image quality assessment database.** Release 2 [http://live.ece.utexas.edu/research/quality/subjective.htm].
- BID-blurred image database. [http://www.lps.ufrj.br/profs/eduardo/ImageDatabase.htm].
- Wang Z, Wu G, Sheikh HR, Simoncelli EP, Yang EH, Bovik AC: **Quality-aware images.** *IEEE Trans Image Process* 2006, **15**(5):1680-1689.

doi:10.1186/1687-5281-2011-3

**Cite this article as:** Chen and Bovik: No-reference image blur assessment using multiscale gradient. *EURASIP Journal on Image and Video Processing* 2011 **2011**:3.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)